

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-339288

(43)Date of publication of application : 24.12.1996

(51)Int.Cl. G06F 3/16
 G06F 3/14
 G06F 3/14
 G06F 17/30
 G10L 3/00
 G10L 5/02

(21)Application number : 07-147291

(71)Applicant : CANON INC

(22)Date of filing : 14.06.1995

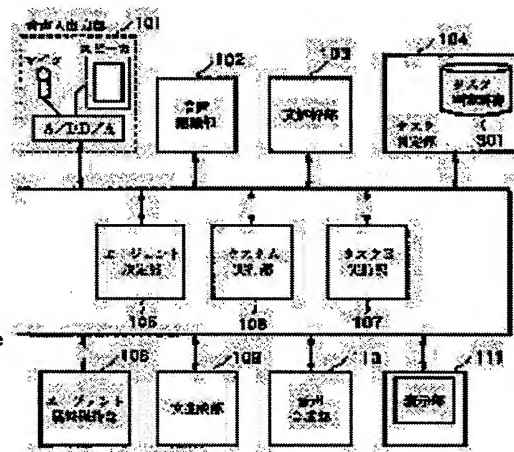
(72)Inventor : SAKAI KEIICHI
 YAGISAWA TSUYOSHI
 FUJITA MINORU

(54) INFORMATION PROCESSOR AND CONTROL METHOD THEREFOR

(57)Abstract:

PURPOSE: To change the form of a response sentence and the synthesized voice for each task and to make a user recognize that the task is changed in a voice interactive processing.

CONSTITUTION: In this information processor (voice interactive device) provided with the plural executable tasks, different agents are allocated corresponding to the respective plural tasks and stored in an agent attribute holding part 108. In the meantime, an input sentence recognized in a voice recognition part 102 is analyzed in an input sentence analysis part 103. A task judgment part 104 decides the task to be used based on the analyzed result and generates a command adapted to the execution task based on the input sentence. A sentence generation part 109 generates the response sentence based on a executed result obtained by executing the command by the execution task. The response sentence is formed corresponding to the attribute information of the agent corresponding to the execution task and a voice synthesis part 110 synthesizes the response sentence corresponding to the attribute information.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
 examiner's decision of rejection or application converted
 registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
 rejection]

[Date of requesting appeal against examiner's decision of
 rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-339288

(43) 公開日 平成8年(1996)12月24日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 3/16	3 3 0	9172-5E	G 0 6 F 3/16	3 3 0 H
3/14	3 3 0		3/14	3 3 0 A
	3 4 0			3 4 0 A
17/30			G 1 0 L 3/00	5 7 1 H
G 1 0 L 3/00	5 7 1		5/02	J
審査請求 未請求 請求項の数12 O L (全 7 頁) 最終頁に続く				

(21) 出願番号 特願平7-147291

(22) 出願日 平成7年(1995)6月14日

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 酒井 桂一

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72) 発明者 八木沢 津義

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(72) 発明者 藤田 稔

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

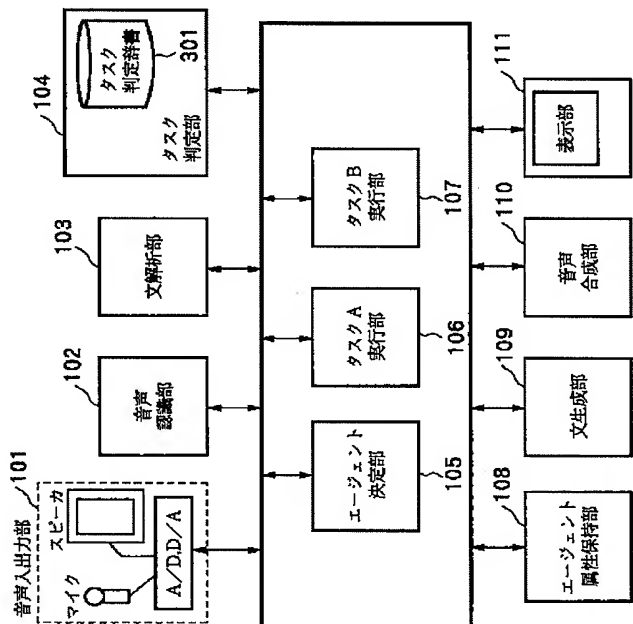
(74) 代理人 弁理士 大塚 康德 (外1名)

(54) 【発明の名称】 情報処理装置及びその制御方法

(57) 【要約】

【目的】 応答文の形態およびその合成音声をタスク毎に変化させることを可能とし、音声対話処理において利用者にタスクが変わったことを認識させる。

【構成】 実行可能な複数のタスクを有する情報処理装置（音声対話装置）において、複数のタスクのそれぞれに対応して異なるエージェントが割り当てられ、エージェント属性保持部108に格納される。一方、音声認識部102で認識された入力文は入力文解析部103で解析される。タスク判定部104はこの解析結果に基づいて使用すべきタスクを決定し、上記入力文に基づく、実行タスクに適応するコマンドを生成する。文生成部109は、このコマンドを実行タスクで実行して得られた実行結果に基づく応答文を生成する。この応答文は、当該実行タスクに対応するエージェントの属性情報に従って形成されるとともに、音声合成部110はこの応答文を上記属性情報に従って音声合成する。



【特許請求の範囲】

【請求項 1】 実行可能な複数のタスクを有する情報処理装置であって、
前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納手段と、
入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定手段と、
前記入力文に基づいて前記決定手段で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行手段と、
前記格納手段に格納されている前記実行タスクに対応する生成情報と前記実行手段で得られた実行結果とに基づいて応答文を生成する生成手段と、
前記生成手段で生成された応答文を出力する出力手段とを備えることを特徴とする情報処理装置。

【請求項 2】 前記生成情報は、前記生成手段で生成される応答文の文体を規定することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】 前記生成情報は、前記生成手段で生成される応答文の表示色を規定することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 4】 前記出力手段は、前記応答文に基づいて音声合成を行い、音声出力を行うことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 5】 実行可能な複数のタスクを有する情報処理装置であって、
前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納手段と、
入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定手段と、
前記入力文に基づいて前記決定手段で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行手段と、
前記格納手段に格納されている前記実行タスクに対応する生成情報と前記実行手段で得られた実行結果とに基づいて応答文を生成する生成手段と、
前記格納手段に格納されている前記実行タスクに対応する生成情報と前記生成手段で生成された応答文とに基づいて音声合成する合成手段とを備えることを特徴とする情報処理装置。

【請求項 6】 入力された音声信号に基づいて入力文を生成する入力文生成手段を更に備えることを特徴とする請求項 5 に記載の情報処理装置。

【請求項 7】 前記生成情報は、前記生成手段で生成される応答文の文体を規定する情報を含むことを特徴とする請求項 5 に記載の情報処理装置。

【請求項 8】 前記応答文を表示する表示手段を更に備えることを特徴とする請求項 5 に記載の情報処理装置。

【請求項 9】 前記生成情報は、前記生成手段で生成される応答文の前記表示手段による表示色を規定すること

を特徴とする請求項 8 に記載の情報処理装置。

【請求項 10】 前記生成情報は、前記合成手段における音声合成処理の各パラメータを含むことを特徴とする請求項 5 に記載の情報処理装置。式。

【請求項 11】 実行可能な複数のタスクを有する情報処理装置の制御方法であって、
前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納工程と、
入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定工程と、
前記入力文に基づいて前記決定工程で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行工程と、
前記格納工程で格納された前記実行タスクに対応する生成情報と前記実行工程で得られた実行結果とに基づいて応答文を生成する生成工程と、
前記生成工程で生成された応答文を出力する出力工程とを備えることを特徴とする制御方法。

【請求項 12】 実行可能な複数のタスクを有する情報処理装置の制御方法であって、
前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納工程と、
入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定工程と、
前記入力文に基づいて前記決定工程で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行工程と、
前記格納工程にて格納された前記実行タスクに対応する生成情報と前記実行工程で得られた実行結果とに基づいて応答文を生成する生成工程と、
前記格納工程にて格納された前記実行タスクに対応する生成情報と前記生成工程で生成された応答文とに基づいて音声合成する合成工程とを備えることを特徴とする制御方法。

【発明の詳細な説明】**【0001】**

【産業上の利用分野】 本発明は複数のタスクを切り替えることで複数種類の処理の実行が可能な情報処理装置及びその制御方法に関し、特に、音声対話装置への適用に好適な情報処理装置及びその制御方法に関するものである。

【0002】

【従来の技術】 近年の自然言語処理技術の向上、音声認識／合成技術の向上により、複数のタスクを扱える音声対話装置が実現されるようになってきた。この種の音声対話装置においては、音声によって入力された情報に基づいて処理すべきタスクを選択、起動し、起動されたタスクにより入力された情報に対応する処理が実行される。そして、この処理結果に基づいて、合成音声が生成され、出力される。

【0003】

【発明が解決しようとする課題】しかしながら、上記従来の装置では、出力である応答文の形態および合成音声が一面的であるため、利用者は、異なるタスクが起動されたか否かをその音声出力から把握することはできなかった。従って、利用者が音声対話装置の各タスクの能力を把握できず、複数のタスクにまたがった処理を要求してしまう可能性がある。即ち、利用者は、音声対話装置の能力を把握できず、タスクの複合を必要とする処理など、当該音声対話装置の能力以上の処理を要求をしてしまうという問題があった。

【0004】本発明は、上記従来技術の問題に鑑みてなされたものであり、タスク毎に出力される応答文の形態を変化させることを可能とし、利用者にタスクが変わったことを認識させることが可能な情報処理装置及びその制御方法を提供することを目的とする。

【0005】また、本発明の他の目的は、応答文の形態およびその合成音声をタスク毎に変化させることを可能とし、音声対話処理において利用者にタスクが変わったことを認識させる情報処理装置及びその制御方法を提供することにある。

【0006】また、利用者にタスクの切り替わりを認識させることで、各タスクの能力を把握させ、これにより、利用者が音声対話装置の能力以上の要求を行うことを抑制する。

【0007】

【課題を解決するための手段】上記の目的を達成するための本発明の情報処理装置は以下の構成を備えている。即ち、実行可能な複数のタスクを有する情報処理装置であって、前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納手段と、入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定手段と、前記入力文に基づいて前記決定手段で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行手段と、前記格納手段に格納されている前記実行タスクに対応する生成情報と前記実行手段で得られた実行結果とに基づいて応答文を生成する生成手段と、前記生成手段で生成された応答文を出力する出力手段とを備える。

【0008】また、好ましくは、前記生成情報は、前記生成手段で生成される応答文の文体を規定する。例えば、「です・ます調」、「である調」、「ございます調」というように文体を変更することで、利用者は容易にタスクの切り換えを認識できようになる。この場合、応答文を音声出力しても表示出力しても利用者はタスクの切り換えを認識できる。

【0009】また、好ましくは、前記生成情報は、前記生成手段で生成される応答文の表示色を規定する。表示色の変化でタスクの切り換えを表すので、利用者は極めて容易にタスクの切り換えを判断できる。

【0010】また、好ましくは、前記出力手段は、前記応答文に基づいて音声合成を行い、音声出力を行う。

【0011】また、上記の目的を達成する本発明の他の構成による情報処理装置は、実行可能な複数のタスクを有する情報処理装置であって、前記複数のタスクのそれぞれに対応して異なる生成情報を格納する格納手段と、入力文を解析して前記複数のタスクの中から使用すべき実行タスクを決定する決定手段と、前記入力文に基づいて前記決定手段で決定された実行タスクに適応するコマンドを生成し、該コマンドを該実行タスクで実行して実行結果を得る実行手段と、前記格納手段に格納されている前記実行タスクに対応する生成情報と前記実行手段で得られた実行結果とに基づいて応答文を生成する生成手段と、前記格納手段に格納されている前記実行タスクに対応する生成情報と前記生成手段で生成された応答文とに基づいて音声合成する合成手段とを備える。

【0012】また、好ましくは、入力された音声信号に基づいて入力文を生成する入力文生成手段を更に備える。音声によって入力文を入力でき、これに対する応答文を音声合成する音声対話装置を構成できるからである。

【0013】また、好ましくは、前記応答文を表示する表示手段を更に備える。応答文を視覚的に認識でき、タスクの切り換えをより容易に把握できる。

【0014】また、好ましくは、前記生成情報は、前記生成手段で生成される応答文の前記表示手段による表示色を規定する。応答文の表示色で切り換えを認識でき、より直観的にタスクの切り換えを判断できるからである。

【0015】また、好ましくは、前記生成情報は、前記合成手段における音声合成処理の各パラメータを含む。

【0016】

【作用】上記の構成によれば、実行可能な複数のタスクを有する情報処理装置において、複数のタスクのそれぞれに対応して異なる生成情報が格納される。入力文を解析することにより、複数のタスクの中から使用すべきタスクが実行タスクとして決定される。そして、入力文に基づく、実行タスクに適応するコマンドが生成される。このコマンドを実行タスクで実行することにより、当該コマンドに対する実行結果（入力文に対応する応答に相当する）を得る。そして、予め格納されている生成情報の中から、実行タスクに対応する生成情報を得て、この生成情報と上記実行結果とに基づいて応答文を生成し、出力する。

【0017】また、上記の他の構成によれば、実行可能な複数のタスクを有する情報処理装置において、複数のタスクのそれぞれに対応して異なる生成情報が格納される。入力文を解析することにより、複数のタスクの中から使用すべきタスクが実行タスクとして決定される。そして、上記入力文に基づく、実行タスクに適応するコマ

ンドが生成される。このコマンドを実行タスクで実行することによりして実行結果を得る。次に、予め格納されている生成情報の中から実行タスクに対応する生成情報を得て、この生成情報と実行結果とに基づいて応答文を生成する生成し、更に生成情報と応答文に基づいて音声合成を行う。

【0018】

【実施例】以下、添付の図面を参照して本発明の実施例を詳細に説明する。

【0019】図1は実施例の音声対話装置のブロック構成を示すブロック図である。同図において、101は音声入出力部であり、マイクより入力された音声をA/D変換して当該装置内部に取り込む一方、生成された応答をD/A変換して、音声としてスピーカより出力する。102は音声認識部であり、音声入出力部101より入力された音声に対して音声認識を行い、入力文を生成する。103は文解析部であり、音声認識部102において認識された入力文に対し、形態素解析、構文解析、意味解析を行う。

【0020】104はタスク判定部であり、文解析部103における入力文の解析結果に基づいて実行するタスクを判別し、実行するタスクに適応したコマンドを生成する。タスク判別部104は、タスクを判別するためのタスク判別辞書301を保持している。本実施例では、説明をわかりやすくするため2つのタスクを扱うものとし、タスクAを内線案内のデータベース検索（例えば、SQLを用いて検索の行えるリレーショナルデータベース）、タスクBを観光案内の全文検索（例えばプレーンテキスト（べた文字列）から条件とした文字列を含む文書を検索するもの）とする。もちろん上記のデータベースの形態は一例であることは、いうまでもない。

【0021】図3はタスク判別辞書301のデータ構成例を示す図である。タスク判別部104では、文解析部103で解析された解析結果中にタスク判別辞書301に登録されている語が現れた場合に、そのタスクを扱うものとして、タスクを切り替える。例えば、「総務部の鈴木部長の内線を知りたい。」という入力文についてタスク判定辞書301を参照すると、「総務部」、「部長」、「内線」という語が存在するタスクAが選択される。

【0022】この結果、タスク判定部104は、タスクAが解釈可能な、select (内線) from (内線テーブル) where (名称=鈴木、所属=総務部、役職=部長) というコマンドを生成し、これをタスクAに渡すことになる。

【0023】同様に、「箱根町にある美術館を教えてください。」という入力文に対しては、タスク判定辞書301の参照によりタスクBが選択される。そして、タスク判定部104はタスクBが解釈可能な、((C分類 美術館) (C所在地 箱根町) (R 名称?)) というコマ

ンドを生成し、これをタスクBに送ることになる。

【0024】105はエージェント決定部であり、タスク判別部104で判別されたタスクに応じてエージェントを決定し、エージェントの切り替えを行う。106、107は、それぞれ異なるタスクを実行するタスクA実行部、タスクB実行部である。本例では、タスクA（内線案内）に対しては「内線エージェント」が、タスクB（観光案内）に対しては「観光エージェント」が夫々選択されるとする。

【0025】108はエージェント属性保持部であり、性別、語調などの文生成情報、及び声の高さ、発声速度などの合成音声情報をエージェントごとに属性として保持するエージェント属性保持部である。図4はエージェント属性保持部108におけるデータ格納状態を表す図である。図4に示されているように、エージェント属性保持部108には、エージェント決定部で決定される「内線エージェント」と「観光エージェント」の夫々に、文生成情報及び合成音声情報が格納されている。

【0026】109は文生成部であり、エージェント決定部105で決定されたエージェントの文生成情報をエージェント属性保持部108より獲得し、これを参照してタスクA実行部106或はタスクB実行部107の実行結果に基づく応答文を生成する。図5は応答文の生成例を表す図である。図5では、上述のコマンドを用いた対話例が示されている。即ち、コマンドに含まれている検索条件「鈴木」、「部長」、「内線」と、タスクAの実行結果として得られた内線番号（ここでは「1234」）を用いるとともに、内線エージェントの文生成情報（女性、ございます調）に基づいて応答文が生成される。もちろん、観光エージェントのように結果のみで応答するようにしても良い。この場合、応答文は、「1234でございます」となる。なお、タスクBの実行結果を出力する観光エージェントによる応答文も、上記内線エージェントと同様である。

【0027】110は音声合成部であり、エージェント決定部105で決定されたエージェントの合成音声情報をエージェント属性保持部108より獲得し、これを参照して、文生成部109で生成された応答文の合成音声を生成する。音声合成部110で生成された合成音声は音声入出力部101のスピーカより音声として出力される。111は表示部であり、音声認識結果、対話の流れ、タスクの実行結果、応答文などを出力する。上記において、女性、男性のエージェント属性は、文生成部109では、応答文において女性の話し言葉、男性の話し言葉のいずれを用いるかを設定するのに参照され、音声合成部110では女性、男性の声を生成するための音素データの設定に用いられる。

【0028】次に図2のフローチャートを参照して、本装置の動作を説明する。図2は実施例の音声対話装置の制御手順を表すフローチャートである。

【0029】まず、ステップS201では、音声入出力部101よりの音声入力待ち、音声入力が行われると、この入力情報を音声認識部102に送る。ステップS202では、音声認識部102により音声認識を行い入力文を得る。また、この認識結果（入力文）を表示部111に表示してステップS203に移る。ステップS203では、文解析部103が、認識結果である入力部に対して形態素解析、構文解析、意味解析を行い、ステップS204に移る。ステップS204では、タスク判別部104において、文解析部103で解析された解析結果から実行するタスクを判別し、実行するタスクに対応したコマンドを生成する。

【0030】ステップS205では、エージェント決定部105において、タスク判別部104で判別されたタスクに応じてエージェントを決定し、切り替える。ステップS206では、タスク判別部104で判別されたタスク実行部（タスクA実行部106あるいはタスクB実行部107のいずれか）において、ステップS204で生成されたコマンドを実行し、ステップS207に移る。

【0031】ステップS207では、文生成部109が、タスクA実行部106あるいはタスクB実行部107で実行された実行結果から応答文を生成する。この応答文の生成にあたって、文生成部109はエージェント決定部105で決定されたエージェントの文生成情報をエージェント属性保持部108より獲得し、この文生成情報を参照しながら応答文を生成する。生成された応答文は表示部111に表示されるとともに、音声合成部110に送られる。

【0032】ステップS208では、音声合成部110が、文生成部109で生成された応答文の合成音声を生じる。この合成音声の生成にあたって、音声合成部110はエージェント決定部105で決定されたエージェントの合成音声情報を属性保持部108より獲得し、この合成音声情報を参照する。生成された合成音声は音声入出力部101でD/A変換されてスピーカより出力される。その後、処理はステップS201へ戻る。

【0033】以上述べたように、上記実施例の音声対話装置によれば、性別、語調、声の高さ、発声速度などの属性を与えたエージェント（担当者）をタスクごとに設定し、タスクの切り替わりによってエージェントを切り替えることが可能となる。すなわち、性別や語調（文生成情報）に基づいて応答文を切り替え、声の高さや発声速度（合成音声情報）に基づいて合成音声の音声切り替えることで、ユーザにタスクが変わったことを意識させることができる。

【0034】従ってユーザは1つのタスクによる処理範囲を把握することが可能となり、これから与えようとする要求が複数のタスクを必要とするような当該音声対話装置の能力を超えたものであるか否かを判断することが

可能となる。

【0035】なお、上記実施例では、説明を簡潔にするために、扱うタスクを2つとして説明しているが、もちろん、3つ以上のタスクを扱うものにも適応できるものである。

【0036】また、上記実施例では、エージェントごとに切り替える応答文の属性として、性別、語調を用いたがこれに限らない。応答文の属性として、例えば表示部に表示する応答文の色や字体を用いることで、エージェントの交代を表すようにしても構わない。

【0037】また、上記実施例では、エージェントごとに切り替える合成音声の属性として、声の高さ、発声速度を用いたこれに限らない。他の例えば、学習音素データや自然降下成分、ポーズ長などを用いることで、エージェントの交代を表すようにしても構わない。

【0038】また、上記実施例では、単一の文生成部に対して属性を与えることによって異なる応答文を生成しているが、生成方式が単なる複数の文生成部を用意し、エージェントごとに文生成部を切り替えることにより、異なる応答文を生成するようにしても構わない。

【0039】また、上記実施例では、単一の音声合成部に対して属性を与えることによって異なる合成音声を生じているが、生成方式が異なる複数の音声合成部を用意し、エージェントごとに合成音声部を切り替えることにより、異なる合成音声を生じるようにしても構わない。

【0040】また、上記実施例では、タスクの切り換えに伴って応答文の形態と合成音声の形態の両方を変化させているが、何れか一方の形態を変化させるようにしても良い。

【0041】以上説明したように、本実施例によれば、扱うタスクを複数持ち、入力された音声認識、文解析し、その解析結果からタスクを判定し、タスクごとに属性を切り替えるエージェントを設定することで、応答文および合成音声の切り替えを実現できる。

【0042】尚、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器から成る装置に適用しても良い。また、本発明はシステム或は装置にプログラムを供給することによって達成される場合にも適用できることはいうまでもない。

【0043】

【発明の効果】以上説明したように本発明によれば、タスク毎に出力される応答文の形態を変化させることが可能となり、利用者にタスクが変わったことを認識させることが可能となる。

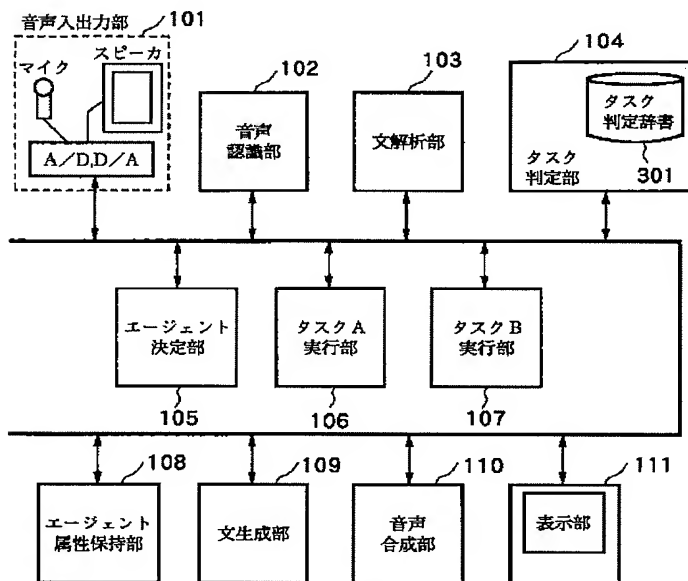
【0044】また、本発明によれば、応答文の形態およびその合成音声のタスク毎に変化させることが可能となり、音声対話処理において利用者にタスクが変わったことを認識させることが可能となる。

【0045】以上のように、利用者にタスクの切り替わ

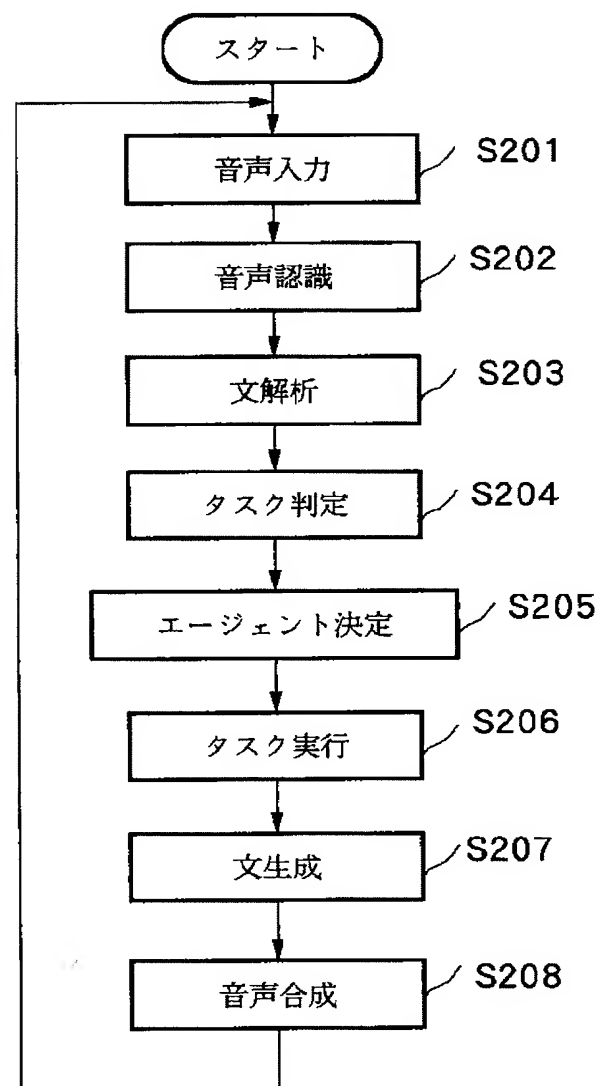
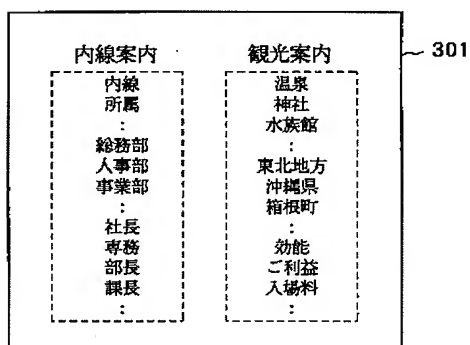
【図4】エージェント属性保持部108におけるデータ格納状態を表す図である。

1 0 1	音声入出力部
1 0 2	音声認識部
1 0 3	文解析部
1 0 4	タスク判別部
1 0 5	エージェント決定部
1 0 6	タスク A 実行部
1 0 7	タスク B 実行部
1 0 8	エージェント属性保持部
1 0 9	文生成部
1 1 0	音声合成部
1 1 1	表示部

【图 2】



【図 3】



【図 4】

108

内線 エージェント		観光 エージェント	
性別	女性	性別	男性
語調	ございます調	語調	ですます調
声の高さ	880	声の高さ	660
発声速度	4	発声速度	3

【図 5】

ユーザ	総務部の鈴木部長の内線を知りたい。
内線エージェント	鈴木部長の内線は 1234 でございます。
ユーザ	箱根町にある美術館を教えてください。
観光エージェント	彫刻の森美術館,成川美術館,箱根仙石原美術館,箱根美術館です。

フロントページの続き

(51) Int. Cl.⁶
G 1 0 L 5/02

識別記号
9289-5 L

庁内整理番号
F I
G 0 6 F 15/403

技術表示箇所
3 1 0 Z